

MTA SZTAKI

Poligén öröklődés modelljei és algoritmusai

Bene Béla

Bevezetés

Közel nyolc éve, hogy megjelentek az első e témával foglalkozó cikkek a genetika irodalmában. Ezek nyomán egyre nagyobb jelentőséget tulajdonítanak napjaink genetikai kutatásai és tanácsadásai a poligén jellegű tulajdonságoknak, rendellenességeknek.

Azóta már számos dolgot tisztáztak és a gyakorlatban eredményesen alkalmaztak, de ezek felvetettek megannyi további problémát is. Ma már nyilvánvaló, hogy magának a poligén rendszernek működése nem kezelhető mechanikusan, nincs a rendellenességek öröklődésére vonatkozó univerzális modell, hanem a modelleket kell úgy kialakítani, hogy azok a lényegtelen effektusok szükséges elhanyagolásával maximálisan figyelembevegyék a tulajdonságot befolyásoló genetikai és konkrét környezeti tényezőket, valamint magyarázatot adjanak az esetleges tünetcsoport háttérstruktúrájára.

Modell

Általában egy tulajdonságot poligénnek nevezünk, ha

I. genetikai komponense több gén együttes hatásából tevődik össze,

II. minden más befolyásoló tényezőt a környezet hatásának tulajdoníthatunk.

Gyakran találkozunk az irodalomban a "multifaktorális" megjelöléssel, amely nem utal az I. teljesülésére, sőt az "egy gén + környezet" modellek (5) esetén teljesen más értelmet takar.

A küszöb-modellek olyan poligén tulajdonságokra vonatkoznak amelyekre a következők is teljesülnek:

III. a tulajdonság fenotipusos megjelenése kvalitatív jellegű /általában nincs konkrét metrikus paraméter, mellyel ha közvetve is, de következtetni lehetne a veszélyeztetettség mértékére/

IV. a fenotipushoz hozzárendelhetünk egy x , $N(0,1)$ eloszlású valószínűségi változót /hajlamot/ mely az additív g , $N(0,h^2)$ eloszlású genetikai ill. a k , $N(0,1-h^2)$ eloszlású környezeti hatásokat magukban foglaló valószínűségi változók összegeként írható fel, feltéve még g és k függetlenségét.

$$x = g + k \quad |1|$$

Ha a manifesztálódott tulajdonság populációbeli valószínűségét p jelöli, úgy definiálhatunk egy olyan K küszöböt, melyre

$$\Phi(K) = 1 - p \quad |2|$$

teljesül, ahol Φ a standard normális eloszlásfüggvényt jelöli.

Ahhoz, hogy modellünk az öröklésmenetre ill. meghatározott utód veszélyeztetettségére válaszolhasson, feltesszük hogy 5./ n egyedből álló családban az együttes eloszlás n -dimenziós együttes normális eloszlás, melynek paraméterei

az egyedek közötti genetikai r_{ij} / $i, j=1, 2, \dots, n$ / korrelációkból ill. h_i^2 örökölhetőségi értékekből adódnak.

Sajnos, az alkalmazásoknál általában az egyes feltételeket nem ellenőrzik megfelelően, hanem a kvalitatív vagy poligén tulajdonságokból kiindulva használják az egész modellt.

Algoritmusok

Legtöbbször a feltett kérdés az, hogy konkrét utódban milyen valószínűséggel manifesztálódik a tulajdonság, ha családjának állapotát ismerjük. Ennek meghatározása a

$$P(x_{\text{utód}} > K \mid \text{állapota}) \quad \text{rokonság} \quad |3|$$

feltételes valószínűség kiszámításával történik.

Ennek közelítésére az irodalomban két jelentős próbálkozással találkozunk.

C. Smith (4) |3| meghatározásánál az együttes eloszlásokat úgy közelíti, hogy

a/ numerikus integrálást csak a két szülő esetében végez a genotípusos eloszláson végighaladva

b/ az összes szereplő többi rokont pedig egymástól függetlennek feltéve a b_{rf}, b_{rm} regressziós együtthatókon keresztül meghatározott

$$\frac{T_r - (b_{rf} g_f + b_{rm} g_m)}{(1 - R^2 h_r^2)^{1/2}} \quad |4|$$

küszöbtávolsággal veszi figyelembe, ahol g_f, g_m az apa ill. anya genotípusos értéke, T_r az r rokon küszöbértéke, R utód-

rokon korrelációs együtthatója és h_r^2 a megfelelő örökölhetőségi érték.

R. N. Curnow (6) algoritmusában egzaktságra törekedve a feltételes valószínűség kiszámítását igyekszik visszavezetni egyváltozós integrálokból álló kifejezések hányadosainak kiszámítására. Kiindulásnál az

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \varphi(\underline{x}, R) B_1(x_1) \dots B_n(x_n) dx_1 \dots dx_n$$

integrált választja, ahol (\underline{x}, R) a már említett sűrűségfüggvény, $B_i(x_i)$, $i=1, 2, \dots, n$; pedig az i -edik rokon konkrét betegség-állapotától függően $\varphi(x_i)$ -t ill. $1-\varphi(x_i)$ -t jelentik. Algoritmus mintegy három-négy egyedből álló család esetén programozható.

Smith és Curnow algoritmusainak gyenge pontjai indítottak arra, hogy más módon próbáljuk a közelítést. Több numerikus módszer is kínálkozik, de reális feladat megoldására alkalmas eloszlásfüggvény esetén a numerikus integrálás ill. szimulációs módszerek alkalmatlanok.

Kiindulva a Taylor-sorok jó tulajdonságaiból, elkészítettünk egy algoritmust és ennek programját, melyben legfeljebb 10 dimenziós standard normális eloszlás sűrűségfüggvényét a 0 helyen a nyolcadik tagig Taylor-sorba fejtettünk, majd tetszőleges határok között integráltuk.

$$\int_{A_1}^{F_1} \dots \int_{A_n}^{F_n} \varphi(\underline{x}) dx_1 \dots dx_n \sim \int_{A_1}^{F_1} \dots \int_{A_n}^{F_n} \sum_{t=0}^k \frac{\varphi^{(t)}(Q)(\underline{x}) t}{t!} dx_1 \dots dx_n$$

Sikerült az egyes deriváltak előállítására egy rekurzív összefüggést megadni. Az első eredmények szerint igen lassan konvergáló sort ad a sorfejtés. Ezért sortranszformációk alkalmazásával próbáltuk a konvergenciasebességet növelni. Az u.n. Euler-féle

$$\sum_{t=0}^{\infty} (-1)^t a_t = \sum_{t=0}^{\infty} (-1)^t \frac{\Delta^t a_0}{2^{t+1}} \quad |7|$$

sortranszformáció gyengének bizonyult. Ennek többszöri egymásutáni alkalmazásának transzformációs formuláját levezettük,

$$\sum_{t=0}^{\infty} (-1)^t a_t = \sum_{t=0}^{\infty} (-1)^t \left(\frac{1}{z+1}\right)^{t+1} a_t \sum_{s=0}^{\infty} \binom{s+t}{s} \left(\frac{z}{z+1}\right)^s \quad |8|$$

majd N-edik szeletig menve, bizonyos átalakítások után, folytonos paraméterre kiterjeszthető transzformációt kaptunk.

$$\left(\frac{z}{z+1}\right)^{N+1} \sum_{t=0}^N (-1)^t a_t \sum_{s=t+1}^{N+1} \binom{N+1}{s} \frac{1}{z^s} \quad Z=0 \quad |9|$$

Z megfelelő értékével, melyet egyéb dolgok megszabnak a jelenlegi tesztelések során harmadik-negyedik tizedesben találhatók az eltérések.

További kutatási irányok

a/ endogám népességben mennyire módosulnak az általános törvényszerűségek

b/ szűk rokonsági körben, különösen testvérek és iker-testvérek esetén a környezeti hatások nyilván nem korrelálatlanul jelentkeznek az egyedekben; hogyan kell ebben az esetben modellünket módosítani

c/ több olyan rendellenességet ismerünk, ahol a küszöb-

jelleg mellett dominanciahatást is észlelhetünk; itt vagy új modellben kell dolgozni, vagy bizonyos feltételes mezőkön megtartható küszöbmodellben

d/ metrikus paraméterekkel is kapcsolatba hozható poligén rendellenesség esetén a küszöb állandóságának eldöntése valamint a metrikus skálán való becslése.

I r o d a l o m

- (1) Falconer, D.S. The inheritance of liability to certain diseases, estimated from the incidence among relatives. Ann. Hum. Gen. 29. 51-76. /1965/
- (2) Falconer, D.S. The inheritance of liability to disease with variable age of onset, with particular reference to diabetes mellitus. Ann. Hum. Gen. 31. 1-20 /1967/
- (3) Smith, C. Heritability of liability and concordance in monozygous twins. Ann. Hum. Gen. 34. 85-91. /1970/
- (4) Smith, C. Recurrence risks for multifactorial inheritance. Ann. J. Hum. Gen. 23. 578-588. /1971/
- (5) Reich, I. James, W.J. and Morris C.A. The use of multiple thresholds in determining the mode of transmission of semi-continuous traits. Ann. Hum. Gen. 36. 163-184 /1973/
- (6) Curnow, R.N. The multifactorial model for the inheritance of liability to disease and its implications for relatives at risk. Biometrics, 28. 931-946. /1972/